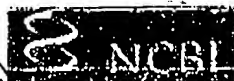


09/17/55



PuroMed

FOUR

BLAST

CMM

Taxonomy

Structure

Match: 5 Mismatch: 2 gap open: 1 gap extension: 2
x dropoff: 100 expect: 10.000 wordsize: 7 Filter ☒ Align

Sequence 1 gi 4538106 Sequence 3 from Patent WO9740167. Length 8209 (1..8209)

Sequence 2 gi 9630311 Gibbon leukemia virus, complete genome Length 8088 (1..8088)

-----The statistics (hitscore and expect value) is calculated based on the size of nr database

----- If protein translation is reversed, please repeat the search with reverse strand of the query sequence

Score = 1583 bits (21250), Expect = 0.0
Identities = 5776/8943 (64%), Gaps = 1625/8943 (18%)
Strand = Plus / Plus

Query: 2 tgggtgtacgactgtg-ggcccccag--cgcgct-tg--gaataaaaaatcctcttgcttgcctt
| | | | | | | | | | | | | | | | | | | | | |
sp1ctf: 12 taqaq-a-qactgtagccgccgqqtaccgc-tgtgtccaataaaa--cctcttgctgat

```

Query:      56  gcatcaaga-coget--tctcgtgagtgat--ttg--ggggtgcgcct--ct-tccga-
              ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Sbjct:      67  qaarcoqgaagccg-ttgtatcgt---tgtctctggggagggttctctctaactat-tgac

```

```

Query:      104  -gcccggaac-gaggggggattgt-tcttttactggcccttccatttgggtgcgttggg-cgggg
           ||||| || ||||| || ||||| ||||| ||||| ||||| ||||| |||||
Sbjct:     122  cgccc--acttcggggg--tctcacatttggggggc---tc---gtccc--ggatcga

```

```

Query:      160 aa----atcctgcgaccacc--ccttacacc--c--gag--aacccgactrgg--ag-gta
              ||      ||      ||      ||      ||      ||      ||      ||      ||      ||
Sbjct:      160 aacccccaccagggaccaccgacc--caccaaaggggaggttaa--g-c-tggccagcg-a

```

```

Query:      205  aag--g-g-atc-ccttt-tg-g---aa---chnnnnnnnnnnnnncc-ggcc--g-gcgc--
           | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Sbjct:      220  ccgttggtgtgtctcgccttcgtgtcctaagtc-cgtaattctgacgtccttgtgtgtctc

```

```

Query:      247  ---tctgt-tctga---gt--gtctg--t---tttcgggtga-tgcgcgc---tttcgg
              ||||| |||||  || |||||  || ||||| ||||| |||||
Sbjct:      279  gcttctgtgtctgaagaccgtaacctctgactgccctc--gtaagtgcgcgcatttttttg

```

```

Query:                287   ttctgcag-ctgtctctctcagaccgt-aagga--ctg-gag-gac-tgtga-t--c-agca
                        ||| | | | | | | | | | | | | | | | | | | | | |
sbjct:                337   tt-cagctctgt--tc-cga---gtgaatcactctgcgaagtgcagctgttgagtagcggaac

```

```

Query:      336  gacgctgctaggaggatcacaggct-gccaccctggggggacgccccggagggt---ggggg-
             |||||  |||||  ||  ||  ||  ||  ||||  ||||  ||  ||||
Sbjct:      390  gacgtgttcgg-ggctcaccgcctggtaatccaggagacgtcccaggga--tcaggggag

```

<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?1>

21/02/01

Query: 392 gagccagggaagcctggtggtctcctactgtcggtcagaggacagagttctgttctgaa
Sbjct: 447 ga-ccagggaagcctggtggtac-cc--c--tcggt-a-a---cg-ggtc-gttg-tgac

Query: 452 gcgaaagcttc-cccctccgcg---g-c-cg-tccgactct-tttgocctgct-tgtggaa
Sbjct: 492 ccgatttcctcgcgcgtctg-gtaagacgogctctgaatctgatt--ct-ctctctcg--

Query: 503 gagcgggacgggtcg-cg-tgtgtctgg-atctgttgggttctgtctcgtg---tgtctt
Sbjct: 546 gtccg---c---tcgcccgcgtctctggtttcttttctgttc-gtttc-tggaaagcctc

Query: 557 tgtcttgt-gc-gtcctctgtctac--agt-tttaatatgggac-agacagtgaactacccc
VIRAL CORE PROTEIN 1 M G Q T V T T P
Sbjct: 598 tg---tgtcacagtctttctctcccaaatcatcaatatgggacaagataat-tctacccc
gag polyprotein 1 M G Q D N S T P

Query: 611 ccttagt-t---tgactctcgaccattggactgaagttagatccagggctcataattgt
VIRAL CORE PROTEIN 9 L S L T L D H W T E V R S R A H N L
Sbjct: 654 ---ta-tctccctcactctaaatcactggagagatctgagaacaaggctcacaatctat
gag polyprotein 9 I S L T L N H W R D V R T R A H N L

Query: 667 cagttcaggttaagaagggaaccttggcagacattctgtgctctgaatggccaacattcg
VIRAL CORE PROTEIN 27 S V Q V K K G P W Q T F C A S E W P T F
Sbjct: 710 ccgtggaaatcaaaaagggaataatggcagactttctgttccctccgagtggtccacattcg
gag polyprotein 27 S V E I K K G K W Q T F C S S E W P T F

Query: 727 atgttgatggccatcagagggaacctttat---tctgaaattatccctgggtgtaagg
VIRAL CORE PROTEIN 47 D V G W P S E G T F N S E I I L A V K
Sbjct: 770 gagtggttggtggcggggagggaacttttaattctctgtctatt-t--ttgcagttaa--
gag polyprotein 47 G V G W P P E G T F N L S V I F A V K

Query: 784 caatcatt---tttca-gactggaccgggtctctcctgatcaggag--ccctatatcc
VIRAL CORE PROTEIN 66 A I I F Q T G P G S H P D Q E P Y I
Sbjct: 825 -aaagattgtctttcagga---ga-acgggggacatccggacca--agttccatatatcg
gag polyprotein 66 K I V F Q R N G G H P D Q V P Y I

Query: 838 t--tacgtggcaagatttgg--cagaagatccctccgccaatgggttaaacatgggttaaat
VIRAL CORE PROTEIN 84 L T W Q D L A E D P P P W V K P W I N
Sbjct: 878 tggta--tggcaggacctcggccaga--atccccaccatgggt---gccca-gcct---
gag polyprotein 83 V V W Q D L A Q N P P P W V P A

Query: 894 aaaccaagaagac-aggtccccgaatcctggctcttggaga---gaaaaacaacact-
VIRAL CORE PROTEIN 103 K P R K P C P R I L A L G E K N K H
Sbjct: 926 ---cc-----gccaaaggtcgtg--t--tgtctct--gatacccg-aagac--cagtt
gag polyprotein 99 S A K V A V V S D T R K P V

Query: 949 -cggccga-aaaagtcgagc-cct---cttctctgtatctaccccgag-ctc-gaggagc
VIRAL CORE PROTEIN 121 S A E K V E P S S S Y L P R D R G A
Sbjct: 967 gggg-ggaggcca-tc-agtctctcccccagcccccatcta-cccg-gcaacagagca-c
gag polyprotein 113 A G R P S A P P R P P I Y P A T D D

Query: 1001 ---cgccgact-tggccggaaccccaacctgttcccccaaccccttatccagcacagggt
VIRAL CORE PROTEIN 139 A D L A G T P T C S P T P L S S T G
Sbjct: 1021 ttaactcctctct---ctgaaccc-ac-g-cccccgccc--tatccggc---ggc
gag polyprotein 131 L L L L S E P T P P P Y P A A

Blast Result

Page 3 of 16

Query: 1057 gctgtgaggggaacctct-gcccctc---c--t-gga--gc--tccgg-----tg---g-t
VIRAL CORE PROTEIN 157 C G E G T S A P P G A P V V
||| | |||| ||||| | | ||| | | ||| | |
Sbjct: 1065 actgcca--ccccctctggccctcaggcgatcgaccgccgtcaggccagatgccgat
gag polyprotein 146 L P P P L A P Q A I G P P S G Q M P D

Query: 1097 -ggagggga-cct--gctgccgg--gactcggagccggagagggccg--acccggag--cg
VIRAL CORE PROTEIN 171 E G P A A G T R S R R G A T P E R
| | | | | | | | | | | | | | | | | | | | | |
Sbjct: 1123 agtagcgatcctgaggggcccagccg-ctggga-ccag-gagtcgccgtgcccgagtcga
gag polyprotein 166 S S D P E G P A A G T R S R R A R S P

Query: 1148 gacagacgaga-tcgcg--at-attaccgctgcgacactatggccctc---ccat-g--
VIRAL CORE PROTEIN 187 T D E I A I L P L R T Y G P P M
| | | | | | | | | | | | | | | | | | | | | |
Sbjct: 1180 g-cagao-a-actcgggtcctgaot-coactgtg-a-ttttgcctcgcgagccatagga
gag polyprotein 184 A D N S G P D S T V I L P L R A I G

Query: 1197 --ccagg--g-ggccaattgc--agccctc--cagattggcccttttcttctgcagat
VIRAL CORE PROTEIN 204 P G G Q L Q P L Q Y W P F S S A D
| | | | | | | | | | | | | | | | | | | | | |
Sbjct: 1234 ccccgccgagcccaatggcctggcctcctacaaatattggccttttcttctgcagat
gag polyprotein 202 P P A E P N G L V P L Q Y W P P S S A D

Query: 1248 ctctataattggaaaactaaccatccccctttctcggaggatccca--acgcctcagcg
VIRAL CORE PROTEIN 221 L Y N W K T N H P P F S E D P Q R L T
| | | | | | | | | | | | | | | | | | | | | |
Sbjct: 1294 ctttataattggaaatctaatccctcttttctga--aaacccagcaggtctcagcg
gag polyprotein 222 L Y N W K S N H P S F S E N P A G L T

Query: 1306 ggttggtggagtccttatgttctctcaccagcctacttgggatgattgtcaacagctgc
VIRAL CORE PROTEIN 240 G L V E S L M F S H Q P T W D D C Q Q L
| | | | | | | | | | | | | | | | | | | | | |
Sbjct: 1352 ggctccttgagtccttatgttctctcaccagcctacttgggatgattgtcaacagctgc
gag polyprotein 241 G L L E S L M F S H Q P T W D D C Q Q L

Query: 1366 tgcagacactcttcacaaccgaggagcgagagagaattctgttagaggctagaaaaaatg
VIRAL CORE PROTEIN 260 L Q T L F T T E E R E R I L L E A R K N
| | | | | | | | | | | | | | | | | | | | | |
Sbjct: 1412 tacagattctttcaccactgaggaacgggaaagaattctcctggaggcccgcaaaaatg
gag polyprotein 261 L Q T L F T T E E R E R I L L E A R K N

Query: 1426 ttcc-tggggccgacggggcgaccc-acgcag-ttgcaaaa--tgagattgacatg-gg-a
VIRAL CORE PROTEIN 280 V P G A D G R P T Q L Q N E I D M G
| | | | | | | | | | | | | | | | | | | | | |
Sbjct: 1472 -tcttggggacaatggg-gccctacacagctcg-agaacct--catt-a-atgaggcc
gag polyprotein 281 V L G D N G A P T Q L E N L I N E A

Query: 1479 ttcccttgactcg-cc-c-cggttgggactacaacacggctgaaggtagggagag-ct-
VIRAL CORE PROTEIN 298 P P L T R P G W D Y N T A E G R E S
| | | | | | | | | | | | | | | | | | | | | |
Sbjct: 1525 ttccctcacaatcgacctcag---tggtattacaacacagccgaggtaggagagctctt
gag polyprotein 299 F P L N R P H W D Y N T A A G R E R L

Query: 1534 -tgaaaatctatgccagg-ctctggtggcggtctc--cggggc-gcctcaagacggcc
VIRAL CORE PROTEIN 316 L K I Y R Q A L V A G L R G A S R R P
| | | | | | | | | | | | | | | | | | | | | |
Sbjct: 1582 atg---gtctacggcc-ggaotctagtggtgaggtctcaaaagggag-etc---ggcgcc
gag polyprotein 318 L V Y R R T L V A G L K G A A R R P

Query: 1589 cactaatttggctaaggtaagagaggtgatgcagggacc-gaacgaacctccc-tcggtg
VIRAL CORE PROTEIN 335 T N L A K V R E V M Q G P N E P S V
| | | | | | | | | | | | | | | | | | | | | |
Sbjct: 1635 taccatttggctaaggtaagagaggtgatgcagggaccggca-gaac--cccttcggt-
gag polyprotein 336 T N L A K V R E V L Q G P A E P P S V

Query: 1647 tttcttgaga--ggctcatggaagcct-tcaggcggttcacccctcttgat-cttacctc
<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?l> 21/02/01

<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?>

21/02/01

```

Query:      2274  --ccattaggaaaaacta----aaagnnnnnnnnntcctgggtgatgggtgccacagg--gca
REVERSE    38      P L G K L      K E K K S W V M G A T G Q
          ||| ||||| || || || || || || || ||
Sbjct:     2305  ccca-tgggaaaaagtagggctccagcggacggtcgTg---ga-aggagcgacaggcgagca
pol polyprotein 38      P M G K V G S R R T V V      E G A T G S

```

```

Query:                2327 acgg-cagtatccatggactaccggaagaacccgrr--gacttggga-gtgggacgggtaa
REVERSE TRANSCRIPTASE 55  R  Q  Y  P  W  T  T  R  R  T  V  D  L  G  V  G  R  V
                        |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Sbjct:                2360 a-ggtc--taccctcgaccacaaag-acrtttaaaaaart-ggacat-aaacaagtga
pol polyprotein       56  K  V  Y  P  W  T  T  K  R  L  L  K  I  G  H  K  Q  V

```

```

Query:                2383 cccactcgtttcttcggtcat-ccctgagtgcccag-tacccct-tctaggtagagacttac
REVERSE TRANSCRIPTASE 74 T H S F L V I P E C P V P L L G R D L
                        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct:                2414 cccactcctctcctgggtcctacccc-gagtgcctctgct-actctgt-tgggaagggagctctcc
pol polyprotein       74 T H S F L V I P E C P A P L L G R D L

```

```

Query:          2440  tgaccaaga-t-gggaggtcaaat-ttc--tttt--gaac-aaggaagaccagaagtgtc-
REVERSE TRANSCRIPTASE  93  L T K M G A Q I S F E Q G R P E V S
                  | | | | | | | | | | | | | | | | | | | |
Sbjct:          2471  taaccag-actaaag-gcccaga-tccagttttccg--ctgagg--gcccacaggttaaca
pol polypeptin   93  L T K L K A Q I Q F S A E G P Q V T

```

```

Query:      2492  tgtgaataacaaaccatcactgtgttgacc--cccc-aa--ttagatgargaatatcgga
REVERSE TRANSCRIPTASE 111  V N N K P I T V L T L Q L D D E Y R
              || | | || | | | | | | | | | | | | | | | | | | |
Sbjct:      2524  tggggagagac--gccc-t-actatg-tg-cctgggtctctaaacctggaagangaataccga
pol polyprotein 111  W G E R P T M C L V L N L E E E Y R

```

```

Query:      2547 ctatat-tctccccaagtaagcctgatcaagatat--a--cagtcctgggttggagcagt
REVERSE    129  L  Y  S  D  Q  V  K  P  D  Q  D  I      Q  S  W  L  R  Q
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Sbjct:      2578 ctacatgaaaagcc-agt--accctcctc----tatcgaccca-tcctggctccagc-tt
pol polyprotein 129  L  H  E  K  P  V  P  S  S      I  D  P  S  W  L  Q  L

```

```

Query:      2602 ttcccca-agcctgggcagaaaccgcagggatgggtttggcgaagcaagtcccccaag
REVERSE    147 F P Q A W A E T A G M G L A K Q V P P Q
TRANSCRIPTASE  ||||| ||||| ||||| ||||| ||||| ||||| |||||
Sbjct:      2629 ttcccactgtatggcgaagaagagccggcatgggactagccaatcaag-tcccac-cag
pol polyprotein 146 F P T V W A E R A G M G L A N Q V P P

```

```

Query:      2661 -gttattca-a-ctgaaggcca-gtg-ctacaccagtatcag-teagacagtagcccttg
REVERSE TRANSCRIPTASE 167  V I Q L K A S A T P V S V R Q Y P L
              ||| | | | | | | | | | | | | | | | | |
Sbjct:      2687 tggtagt-agagct-aagatcaggtgcct-caccagtggtgtgtc-gacaatatccaatg
pol polyprotein 165  V V V E L R S G A S P V A V R Q Y P M

```

```

Query:      2715 agtagagaggctcgagaagggaatttggccgcgatgttca-aagattaatcc-a-ac--agg
REVERSE TRANSCRIPTASE 105  S R E A R E G I W P H V Q R L I Q Q
              || | | | | | | | | | | | | | | | | | | | |
Sbjct:      2743 agcaagaagctcggaaggtatcagacccacatccagaag-t---tcctagacctagg
pol polyprotein 184  S K E A R E G I R P H I Q K F L D L G

```

```

Query:          2770 gcacccctagt-tcctgtcccaatc-cccttggaatactccc-ctgctaccggttaggaagc
REVERSE TRANSCRIPTASE 203 G I L V P V Q S P W N T P L L P V R K
                  | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Sbjct:          2799 g-gtcttgggtgcctgt-cggtcgccc-tggaataac-ccctotgctaccggraaaaaagc
pol polyprotein  202 V L V P C R S P W N T P L L P V K K

```

```

Query:          2827 ctgggaccaatgattatcgaccagttacaggacttgagagagggtcaataaaaggggtgcagg
REVERSE TRANSCRIPTASE 222 P G T N D Y R P V Q D L R E V N K R V Q
                  | | | | | | | | | | | | | | | | | | | | | | | | | |
Sbjct:          2855 cagggaccaatgactatcgccagttcaagacctgagagaaatataaaaggggtacagg
pol polyprotein 221 P G T N D Y R P V Q D L R E I N K R V Q

```

Query: 2867 acatacaccacacgggtcccggaaccttatca-gctcttgagcgcgoottoogooctgaag
http://www.ncbi.nlm.nih.gov/blast/blast2seq/wblast2.cgi? 21/02/01

Blast Result

Page 6 of 16

REVERSE TRANSCRIPTASE 242 D I H P T V P N P Y N L L S A L P P E R
| | | | | | | | | | | | | | | | | | | | | |
Sbjct: 2916 atattcatccc-acagtcaccaaaccccttacaatcttc-tgagttcccttcgcct-agc-
pol polyprotein 241 D I H P T V P N P Y N L L S S L P P S

Query: 2946 ga-ac-tggtagaaggtatggga-actaaaagatgacttcttctgootgagattaca-cc
REVERSE TRANSCRIPTASE 262 N W Y T V L D L K D A F F C L R L H
| | | | | | | | | | | | | | | | | | | | | |
Sbjct: 2971 tatacttggtagctcagctcttagatc-tcaaggatgccttttctgcctcaggtacatcc
pol polyprotein 260 Y T W Y S V L D L K D A F F C L R L H P

Query: 3001 ccactagccaaccactttttgccttcgaatggagagatccag-gracgggaagaaccggg
REVERSE TRANSCRIPTASE 290 P T S Q P L F A F E W R D P G T G R T G
| | | | | | | | | | | | | | | | | | | | | |
Sbjct: 3030 caac-agccagccgctgttgcggttcgagtggaagaccacagaaaaggtaa-cacaggt
pol polyprotein 280 N S Q P L F A F E W K D F E K G N T G

Query: 3060 cagctcacctggaccgactgcaccaagggttcaagaactc-cccagc-catctttgacg
REVERSE TRANSCRIPTASE 300 Q L T W T R L P Q G F K N S P T I F D
| | | | | | | | | | | | | | | | | | | | | |
Sbjct: 3098 cagctgacctggacgcygctacacagaaggttcaagaactctccc-actc-tcttcgacg
pol polyprotein 299 Q L T W T R L P Q G F K N S P T L F D

Query: 3118 aagccctacacagggacctgggc-caacttcaggatc--caacacccctcaggtgaccctcc
REVERSE TRANSCRIPTASE 319 E A L H R D L A N F R I Q M P Q V T L
| | | | | | | | | | | | | | | | | | | | | |
Sbjct: 3146 aggccctccacagagatttggctc-cttttagggccctcaac-ccc-caggtggtgttac
pol polyprotein 318 E A L H R D L A P F R A L N P Q V V L

Query: 3175 tccagtacgtggatgacctgcttctggcgaggaccacaaaca-g--gactgcttagaag
REVERSE TRANSCRIPTASE 338 L Q Y V D D L L L A G A T K Q D C L E
| | | | | | | | | | | | | | | | | | | | | |
Sbjct: 3203 tccaatatgtggacgacct-ct--tggtggccgccccacatagaaactgcbaaaaag
pol polyprotein 337 L Q Y V D D L L V A A P T Y E D C K K

Query: 3232 gtac---gaaggcactactgctggaatt--gtctgacctaggtacagagcctctgctaa
REVERSE TRANSCRIPTASE 357 G T K A L L L E L S D L G Y R A S A K
| | | | | | | | | | | | | | | | | | | | | |
Sbjct: 3260 gaacacagaa-gctct--tacaggagtttaag--taagttggggcaccgggtacggctaa
pol polyprotein 356 G T Q K L L Q E L S K L G Y R V S A K

Query: 3287 gaaggccagatttcg-caggagagaggttaacata-cttgggggtacagt-ttgc-gggggc
REVERSE TRANSCRIPTASE 376 K A Q I C R R E V T Y L G Y S L R G
| | | | | | | | | | | | | | | | | | | | | |
Sbjct: 3315 gaaggccagctctgcca-gagagaagtccactatc-tgggggtac-ctact-caaggaag
pol polyprotein 375 K A Q L C Q R E V T Y L G Y L L K E

Query: 3343 ggcagcgatgggtga-oggagggaagggaagaactg-tagtcagatccgggccccaa
REVERSE TRANSCRIPTASE 394 G Q R W L T E A R K K T V V Q I P A P T
| | | | | | | | | | | | | | | | | | | | | |
Sbjct: 3371 gaaaaagatggcctaaccna-gccgaaaggctactgtta-tgaaaatccctgttctctac
pol polyprotein 393 G K R W L T P A R K A T V M K I P V P T

Query: 3401 cacagccaaacaagttagagaggtttttggggacagctggatatttcagactgtggatccc
REVERSE TRANSCRIPTASE 414 T A K Q V R E F L G T A G F C R L W I P
| | | | | | | | | | | | | | | | | | | | | |
Sbjct: 3429 gacccccagacaggtccgtgaatttctagggcactgcgggattctgcagggtctgtgaccc
pol polyprotein 413 T F R Q V R E F L G T A G F C R L W I P

Query: 3461 ggggttttgcgaccttagcagc-cccactctaccgctaaccaagaaggggggattct
REVERSE TRANSCRIPTASE 434 G F A T L A A P L Y P L T K E K G G F
| | | | | | | | | | | | | | | | | | | | | |
Sbjct: 3489 tgggtttgcttccctggctgcacc-ttgtaacccctaacaaaaga----gagca-tc-
pol polyprotein 433 G F A S L A A P L Y P L T K E S I

Query: 3520 cc-----tgggctcctgag---caccagaaggcatttgatgc-tatcaaaaaggccctg
REVERSE TRANSCRIPTASE 453 S W A F E M Q K A F D A Y K K A L

<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?>

21/02/01

21/02/01

<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?>

21/02/01

pol polyprotein 843 S A V R E V T S Q C Q A C A M T N A V

```

Query:          4821 aatc-cttcaga-atacctccaggaaagagactaa-gggga-agccaccagggcgt-c
               867  N   P   S   R   I   P   P   G   K   R   L   A   G   S   H   P   S   A
               |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
Sbjct:          4776 caccacctacagagagacc-----ggaaaaaggg-aacgaggagatcgacc-ggcg-tgt
               862  T   T   Y   R   E   T               G   K   R   Q   R   G   D   R   P   G   V

```

```

Query:          4876 actgggaagtggacttcactgaggtaaagcc-ggctaaaracggaatacaatatctattg
REVERSE TRANSCRIPTASE 905 M W E V D F T E V K P A K Y G N K Y L L
                ||||| | | ||||| | | ||||| | | ||||| | | ||||| | | ||||| | | ||||| | |
Sbjct:          4829 actgggaggtagacttcacagaataaagcctgg-tcggtatggaacaagtatctgtta
pol polyprotein 879 Y W E V D F T E I K P G R Y G N K Y L L

```

```

Query:          4936 gtrrrtcttagacaccttt-tcaggatgggttagaggcttatcctactaa---gaaa-gag
REVERSE TRANSCRIPTASE 925  V F V D T F S G W V E A Y P T K      K E
                  || || || | || | || | || | || | || | || | || | || | || |
Sbjct:          4888 gatttcatagata-ctttctcggatgggtagaagctttctaccaaactgaaacg-g
pol polyprotein 899  V F I D T F S G W V E A F P T K T E T

```

```

Query:          4939 acctcaaccgtgggtggcctaagaaataactggaggaaatttttccaagatttggaaacact
REVERSE TRANSCRIPTASE 943  T S T V V A K K I L E E I F P R F G I P
                  | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Sbjct:          4946 cccat-aatcgt-ctg--taaaaaaattattagaagaaattctacccgcgttcgggatccct
pol polyprotein 918  A L I V C K K Y L E S I L P R F G I P

```

```

Query:          5049 aaggtaaatgggtcagacaatggtccag-cttfcggtgccaggttaagtcagggaactggc
REVERSE TRANSCRIPTASE 963 K V I G S D N G P A F V A Q V S Q G L A
                  ||| ||||| ||||| ||| ||| ||||| ||||| ||||| ||||| ||||| |||||
Sbjct:          5002 aaggtaactcggtcagacaatggccccggccttt-gtgtctcagggttaagtcagggaactggc
pol polyprotein 937 K V L G S D N G P A F V A Q V S Q G L A

```

```

Query:          5108 ca-agatattggggattgattggaaacgcattgtgcatacagaccccaagctcaggac
REVERSE TRANSCRIPTASE 983  K I L G I D W K Y H C A Y R P Q S S G
                || ||||| ||||| ||||| ||||| ||||| |||||
Sbjct:          5061 cactca-actggggataaaattggaagttacattgtgcgtatagaccccaagctcaggtc
pol polypeptide 957  T Q L G I N W K L H C A Y R P Q S S S

```

```

Query:          5167 aggtagagagagatgaatagaaaccattaaagagaccctt-accaaat-gaccacagagac
REVERSE TRANSCRIPTASE 1002 Q V E R M N R T I K E T L T K L T T E T
                  ||||| ||||| ||||| ||||| ||||| ||||| |||||
Sbjct:          5120 aggtagaagaatgaacagacaattasagaga-ccttgaccaaatag-ccttagagac
pol polyprotein  976 Q V E R M N R T I K E T L T K L A L E T

```

```

Query:          5225 ---cggcattaatgattggatggctctcctgccttctgtgcttttttagggtaggaacac
REVERSE TRANSCRIPTASE 1022      G I N D W M A L L P F V L F R V R N T
          ||| | | | | | | | | | | | | | | | | | | | | |
Sbjct:          5178 cgggtgg-a--aaagactgggtgacctccttcccttagcgctgcttagggccaggaatac
pol polyprotein  996      G G      K D W V T L L P L A L L R A R N T

```

```

Query:          5282 ccctggcagatttgggctgacccccctatg-aattgctctacggggga----nnnnnnhng
REVERSE TRANSCRIPTASE 1041  P G Q F G L T P Y E L L Y G G      P P P
                |||||
Sbjct:          6235 ccctggcgggttgggttaactccttatgaaatt-ctctatggaggaccacccccctac
pol polyprotein 1015  P G R F G L T P Y E I L Y G G P P P I

```

```

Query:          5337 ttg-g-c---agaaatgaccttgcacatagtgcctgatg-t-gctgctt-tccagcctt
REVERSE TRANSCRIPTASE 1059  L  A  E  I  A  F  A  H  S  A  D  V  L  L  S  Q  P
                ||| | | | |||| | | | | | | | | | |
sbjct:          $294 ttgagtgtggagaaa----ccttg----ggtcgcgatgatagat--ttctccctgtctt
pol polyprotein 1034  L  E  S  G  E      T  L      G  P  D  D  R      F  L  P  V  L

```

```

Query:          5389 -tggt-ctctagggc-tcaaggcgctcg-agtgggtgaggcagc--gagcgtggaagcagc
REVERSE TRANSCRIPTASE 1076  L F S R L K A L E W V R Q R A W K Q
                | | | | | | | | | | | | | | | | | | | |
Sbjct:          5343 at-ttactc-a--cttaaggccttagaaat-tgtaagg-acccaaatc-tgggaccaga
pol polyprotein 1050  F T K L K A L E I V R T Q I W D Q

```

<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi>

21/02/01

Blast Result

```

Query:          5443 tccggggaggcctactcag-gaggagac--ttgcaagtccaca-tcgcttccaagttgga
REVERSE TRANSCRIPTASE 1094 L R E A Y S G G D L Q V P H R F Q V G
                  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct:          5396 tcaaagaggtgta-taagcctgg-taccgtatacaa-tccctcacccg-ttcagggtcdgg
pol polypeptin   1068 I K E V Y K D C T V T Y P H P F Q V G

```

```
Query:          5499 gattc-agt-ctatgttagacgccaccg-tgcaggaaacctcgagactcgggtggaaggga
REVERSE TRANSCRIPTASE 1113 D S V Y V R R H R A G N L E T R W K G
                  || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Sbjct:         5452 ga-tcaagtgtct-tgtcagacgcccatcgacca-gcaccccttgagcctcgggtggaaggc
pol polyprotein 1067 D Q V L V R R H R P S S L E P R W K G
```

```
Query:          5556 ccttatctcgtacct-ttgaccacaccaagcggtgtgaaagtgcgaaggaaatccccactg
REVERSE TRANSCRIPTASE 1132 F Y L V L L L T T P T A V K V D G I P T W
                  ||| || | || | || | || | || | || | || | || | || | || |
Sbjct:         5509 ccataccctggg-gtggctgactaccccgacccgcggtaaaagtcgatggtattggccgtg
pol polyprotein . 1106 F Y L V L L L T T P T A V K V D G I A A W
env protein      1                               M V L L P
```

```

Query:          5615 gatccatgcatccccacgttaagccggcgccacotcccg-actc-g---ggg---tggagag
VIRAL ENVELOPE PROTEIN 1      H H P T L S R R H L P T R G R
REVERSE TRANSCRIPTASE 1152 I H A S H V K P A P P D S G W R
                        | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Sbjct:          5568 ggtccatgctcttcacatcaaacctgcaccaccttcgggcac-cagatgagtgctctgg-gag
env protein     6      G S M L L T S N L H H L R H Q M S P G S
pol polypeptide 1126 V H A S H L K P A P P S A P D E S W E

```

```

Query:          6668  c-cgaaaagaactgagaaatcccccttaagcttcgcctcca-cc-gc--ctggttccttac--
VIRAL ENVELOPE PROTEIN 17   P K R L R I P L S F A S I A W F L T
REVERSE TRANSCRIPTASE 1169 A E K T E N P L K L R L H R L V P Y
                        | | | | | | | | | | | | | | | | | | | | | |
sjbct:          5626  ctggaaaagaactgatcatcccttaagct--gcgt--attcggcggc-gg--cgggacga
env protein     25   W K R L I I L L S C V F G G G G T
pol polypeptide 1145  L E K T D H P L K L R I R R R A P Y

```

```

Query:          5721 -tct--aacaat---aactcc--ccaggccagtagtaaacgcctta-tagacagctc-c-g
VIRAL ENVELOPE PROTEIN 35      L   T   I   T   P   Q   A   S   S   K   R   L   I   D   S   S
REVERSE TRANSCRIPTASE 1187    s   N   N   N   s   P   G   Q   ^^^
                        |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Sbjct:          5679 gtctgcaa-aataagaaacccccaccagccca-t-g---ac-cctcacttggcagggtactg
env protein     43      S   L   Q   N   K   N   P   H   Q   P   M           T   L   T   W   Q   V   L
pol polyprotein 1162    S   A   K   ^^^

```

```

Query:          5770 aacccccca-t--agacert-tatcccttac-ctggctgat-tattgacc--tgatac-g
VIRAL ENVELOPE PROTEIN 51  N P H   R P L S L F W L I I D P D F
                   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct:          5732 -tcccaactggagcgtgtgtctgggatacaaaaggcag-tcca--gcccccttg-gactt
env protein      61  S Q T G D V V W D T K A V Q   P P W F

```

```

Query:          5821 ggtg----tcactgtaaa--t-a-gca---c--tcgagg-t-g---ttg-cttct-aga-
VIRAL ENVELOPE PROTEIN 68  G      V  T  V  N      S      T      R  G      V  A  P  R
                    ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||  ||||
Sbjct:          5787 ggtggccacact-taaacctgatgtatgtgccttggcggctagtccttgagtccttgggat
env protein      79  W  W  P  T  L      K  P  D  V  C  A  L  A  A  S  L  E  S  W  D

```

```

Query:                5860  -----ggcacc--tg-gt-ggc--crgaactgcatttctg-cctcc-gatt--gatta-
VIRAL ENVELOPE PROTEIN 81      S T W W F E L H F C L R L I
                        || || | | | | | | | | | | | | | |
Sbjct:                5846  atcccggaaccgatgtctcgtcctctaacc-g-a-gtcagacctccggactcagactat
env protein            99      I F G G T D V S S S K R V R F F D S D Y

```

```

Query:          5903 ac-cccgtgtgtaaaagca--cacct-----ccca-a-----c--ctagtc-cg--ta
VIRAL ENVELOPE PROTEIN 95  N  P  A  V  K  S      T  P      P  N      L  V  R
                  || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct:          5903 actgcgcg--ctataagcaaatcacctggggagccatagggtgcagctaccctcggggta
env protein     118  T  A  A      Y  K  Q  I  T  W  G  A  I  G  C  S  Y  P  R  A

```

<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?>

21/02/01

```

Query:          6447  taccactggatrgattectaccaa----cagcgctagaa-actcc-ccaggtgttctcgt
VIRAL ENVELOPE PROTEIN 277  T T G L I P T N      T P R N S P G V P V
                  | | | | | | | | | | | | | | | | | | | |
Sbjct:          6484  tcccac-gg--tga-gaaaaacattgttac-ccaa-aacactccgc-----tcc--
env protein     312  P T      V R K T I V T L N T P P      P

```


VIRAL ENVELOPE PROTEIN 499 H A A M T E D L R A L K E S V S N L E
 | | | | | | | | | | | | | | | |
Sbjct: 7140 agat-c-gccataga-tgctgacctcggcgccctcaa-gactcagtcagcaagttagag
env protein 530 Q I A I D A D L R A L Q D S V S K L E

```

Query:          7171 gactccctgacttctttctctgaagtggcctctacagaaccggag-gggattaga-tctgc
VIRAL ENVELOPE PROTEIN 518  E S L T S L S E V V L Q N R R G L D L
                  || || || || || || || || || || || || || || || || || || || || ||
Sbjct:          7196 gactcactgtctccctgtccgaggtagtgtctcctcaaataggagagggcctt-gact-tgc
env protein     549  D S L T S L S E V V L Q N R R G L D L

```

```

Query:                7229  tgttttctaagagaaggtgggttatgtgcagccttaaaagaagaatgtergttctatgtag
VIRAL ENVELOPE PROTEIN 537  L F L R E G G L C A A L K E E C C F Y V
                        |||||
Sbjct:                7254  tgttttctaaagaaggtggcctctgtgcggccctaaaggaagagtgctgtttttacatag
env protein          568  L F L K E G G L C A A L K E E C C F Y I

```

```

Query:          7289 atcactcaggagcca-t-cagagactccatgaacaagct--tagaaaaa--agttagaga
VIRAL ENVELOPE PROTEIN 557 D H S G A I R D S M N K L R K K I E
               ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
Sbjct:          7314 accactcaggtg-cagtac-gggactccatgaaaaaactcaagaaaaaaactggata-aaa
env protein     588 D H S G A V R D S M K K L K E K L D K

```

```

Query:                7343 ggc-g-tcgaaggggaagagaggctgaccaggggtggttgaaggatggttcaacaggtc
VIRAL ENVELOPE PROTEIN 575 R R R R R R E A D Q G W F R G W F N R S
                        | | | | | | | | | | | | | | | | | |
Sbjct:                7371 gacagttag-agcgcagagaag----ccaaaactggtatgaaggatggttcaataactc
env protein           607 R Q L E R Q K S Q N W Y E G W F N N S

```

```

Query:                7401 tctctggatgaccaccctgctttctgc--tctgacggggccctagtagtctctgacctg
VIRAL ENVELOPE PROTEIN 595  P W M T T T L L S A L T G P L V V L L L L
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Sbjct:                7426 ccctgggttcactaccctgctatcaaccatc-g-ctggggccctattactcctctctctg
env protein           626  P W F T T L L S T I A G P L L L L L L L

```

```
Query:          7459 rtacttacagt--t-gggccttgcttaattaataggtttggtgc-ctttggttaga-ga-a
VIRAL ENVELOPE PROTEIN 614   L L T V      G P C L I N R F V A F V R E
                || | | | | | | | | | | | | | | | | | |
Sbjct:          7484 ttgct--ca-tccctgggccatgcacatcaataagtttagt-caattcatca-atgata
env protein     645   L L      I L G P C I I N K L V Q F I N D
```

```

Query:          7613 cgagtgagtgca-gtcaga-tcatggtacttaggcacagtagccaaggccttct-gagc
VIRAL ENVELOPE PROTEIN 632  R V S A V Q I M V L R Q Q Y Q G L L S
                || | | | | | | | | | | | | | | | | | | | | | | | | | |
Sbjct:          7549 gga-taagtgcgtgttaaaattc-tggtccttagacaa-aatctc-aggcc--ctagag-
env protein     663  R I S A C ^^^

```

```

Query:      7570 caaggagaaactgacctctagccttccagttctaagattagaactattaacaagacaag
VIRAL ENVELOPE PROTEIN 651  Q G E T D L ^^^
              ||      ||| ||| ||      ||      ||| ||| ||| ||| ||| ||| |||
Sbjct:      7692 -aa--cgaaggttaacotttaatttt---gctctaagattagagctattcacaagagaa-

```

```

Query:      7630 aagrggggaatgaaagga-tg--aaaatgcaacctaacctccag-aac-ccaggaagt
             | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Sbjct:      7644 atg-ggggaatgaaagaagtgttttttttagcc-aa--ct-gcagtaacgcc--attt

```

```

Query:      7685 taata--aaaagcctctaaatgccccccg--aattacagaccct--g--ctgdcctgccagta
             | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Sbjct:      7696 tgctagggcaca-c-ctaaa-ggataggaaaaatacag---ctaagaacagg--gcc--aa

```

```

Query:      7737 aatagg-ca---gaaggta--cacttctattgttccagggcctgcatactggcctaa
             || ||| |   || |||| ||| || | || || || || || || || ||
Sbjct:      7745 aacaggatatctgt-gtgtcatgcac--c---tggggc---ccggc---cagggc--aa

```

Query: 7791 gta-aga----taacaggaaatgagttgact-aatc-gcttat-tggatt-ctgaa-aa

<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?>

21/02/01

Blast Result

Sbjct: 7790 ggacagaggggtccca-gaaat-agatgagtcac-cagc--a----gtttccagcaagga
Query: 7842 ctgactgg--caccataga-agaattga-ttacacattgacag--cccta--gtgaccta
Sbjct: 7841 caga-ggggttc-ccagaaataga--tgagtca-aca--g-cagtttcc-aggggtgccc--
Query: 7894 tctcaactgcaatctgtcactctgccaggagcccacgcagatgctggacctccggagcta
Sbjct: 7890 -ctcaaccg---t-t-tca-----agga-ctc-c-c-at---ga---ccgg-g---
Query: 7954 ttttaaaatgattggtccacggagcgggctctcgatattttaaaatgattggtcc-a t
Sbjct: 7921 -----aat-t-----cac-----ccctggc-ct---ta-tttgaactaa-----ccaat
Query: 8013 ggagcgc-gggctctcgatattttaaaatgattggttctgacgcacag-gctttgttg-
Sbjct: 7954 -ta-c-cttgctctcgct-----tctgt-ac-c-c-gcgttt-ttg
Query: 8070 tga-accocataa-aagctgtcccgattc-cgactcggggc-cgcagtcctctacccct
Sbjct: 7990 t-ata-----a-ataagct---cagaaactc-caccggagcgc-cagtcct-ta-----
Query: 8126 gcgtg-gtga--cgactgtgggcccagggcggttggaaataaaatcctctgtgttt
Sbjct: 8033 gagagactg-agccgccc-ggtaccc-gtgtg--tccaataaaa--cctctgtgtgt
Query: 8183 gca 8185
Sbjct: 8086 gca 8088

Score = 113 bits (1449), Expect = 9e-22
Identities = 577/1182 (48%), Gaps = 460/1182 (38%)
Strand = Plus / Plus

Query: 7415 ccc---tgetttctg-ctctga--cg---ggg--ccoctagtagtcc--t---gtc--
VIRAL ENVELOPE PROTEIN 599 T L L S A L T G P L V V L L
Sbjct: 2 cgcagtccttagagagactgagccgcccgggtaccggt-gt-gtccaataaaacctctt
Query: 7456 -ctg-ttacttac--ag--ttgggc-c-ttg---ctt---a--at---taa-ta--g
VIRAL ENVELOPE PROTEIN 613 L L L T V G P C L I N R
Sbjct: 60 gctgattgcatcggagccggtgtctcgttctccttgggaggggtttctcctaactattg
Query: 7491 ---g---tt---tgt-tgc-c-tttg---t--t---aga---gaa---cgagtga
VIRAL ENVELOPE PROTEIN 624 F V A P V R E R V
Sbjct: 120 accgcccacttcgggggtct-cacatttgggggtcgtccgggatcggaaacccaccca
Query: 7520 gtg--ca--g-tcca-gatc---atggta--cttaggcaacag-tacc-aag-g---c--
VIRAL ENVELOPE PROTEIN 634 S A V Q I M V L R Q Q Y Q G
Sbjct: 178 gggacccagagcccaacggga-dgtaagct--dgc--cagcgaccgtgtgtgtctc
Query: 7561 -ctctgag-cgaag-gagaaa--atgac-ct-atagcct-tc-c---cagt-tct-a-
VIRAL ENVELOPE PROTEIN 648 L L S Q G E T D L ^^^
Sbjct: 234 gcttctgtgtcctaagtcgtaattctgactgtccttgtgtgtctcgtctctgtgtgag
Query: 7605 a--g-a--t-t-a--gaactat-taa---caagaca-----ag---aagt-----g
Sbjct: 294 aocgtaactctgactgoot-tgtaagtgcgcg-catttttttggtttcactctattccg

<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?1>

21/02/01

<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?1>

21/02/01

Blast Result

Subject: 1095 cggaccgcccgtcaggccagatgcc-c--gatag-tagcgate 1132
gag polypeptide 156 G P P S G Q M P D S S D

CPU time: 6.79 user secs.

0.19 sys. secs

6.98 total secs.

Gapped
Lambda K H
0.0514 0.0120 0.0135

Gapped
Lambda K H
0.0514 0.0120 0.0135

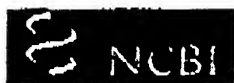
Matrix: blastn matrix:5 -2
Gap Penalties: Existence: 1, Extension: 2
Number of Hits to DB: 7964
Number of Sequences: 0
Number of extensions: 7964
Number of successful extensions: 13
Number of sequences better than 10.0: 1
length of query: 8209
length of database: 2,812,937,656
effective HSP length: 1920
effective length of query: 6289
effective length of database: 2,812,935,736
effective search space: 17690552843704
effective search space used: 17690552843704
T: 0
A: 0
X1: 135 (10.0 bits)
X2: 1347 (99.9 bits)
S1: 250 (24.9 bits)
S2: 463 (40.7 bits)

<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?l>

21/02/01

Blast Result

Attachment B



Blast 2 Sequences results

PubMed

Entrez

BLAST

OMIM

Taxonomy

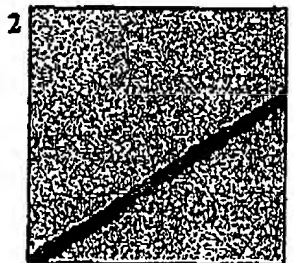
Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTN 2.1.2 [Nov-13-2000]

Match: Mismatch: gap open: gap extension: x_dropoff: expect: wordsize: Filter ☐

Sequence 1 gi 4538106 Sequence 3 from Patent WO9740167. Length 1575 (588 .. 2162)

Sequence 2 gi 9630311 Gibbon leukemia virus, complete genome Length 2370 (631 .. 3000)



NOTE: The statistics (bitscore and expect value) is calculated based on the size of nr database

NOTE: If protein translation is reversed, please repeat the search with reverse strand of the query sequence

Score = 362 bits (4802), Expect = 2e-98
 Identities = 1164/1749 (66%), Gaps = 345/1749 (19%)
 Strand = Plus / Plus

```

Query:          588  atgggac-agacagtgaactaccccccttagt-t---tgactctcgaccattggact
VIRAL CORE PROTEIN 1  M G Q T V T T P L S L T L D H W T
                   ||||| ||| | | ||||| | | | | ||||| | | |||
Sbjct:          631  atgggacaagataat-tctacccc---ta-tctccctcaactotaatcactgga--
gag polyprotein  1   M G Q D N S T P I S L T L N H W

```

```

Query:          642  -gttagatccaggggtcataattgtcagttcagg---ttaagaagggaccttggc
VIRAL CORE PROTEIN 19 V R S R A H N L S V Q V K K G P W
                   ||||| ||||| ||| | | ||||| | | ||||| |||
Sbjct:          684  tgtgagaacaaggggtcacaatctatccgt---ggaaatcaaaaagggaaaatggc
gag polyprotein  18  V R T R A H N L S V E I K K G K W

```

```

Query:          698  tttctgtgocctctgaatggccaacattcgatgttggtggccatcagaggggacct
VIRAL CORE PROTEIN 38 F C A S E W P T F D V G W P S E G T
                   ||||| ||||| ||| | | ||||| | | ||||| |||
Sbjct:          741  tttctgttccctccagtggtggccacattcggcgtgggtggccaccggagggaactt
gag polyprotein  38  F C S S E W P T F G V G W P P E G T

```

```

Query:          758  t---tctgaaattratccgtggtgtaaggcaatcatt---tttca-gactggaccc
VIRAL CORE PROTEIN 57 S E I I L A V K A I I F Q T G P
                   | |||| ||| | | | |||| | | ||| ||||| || |
Sbjct:          801  tctctctgtcatt-t---ttgcagttaa---aaagattgtctttcagga---ga-ac
http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?0 22/02/01

```

<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?0>

22/02/01

Blast Result

Attachment C



Blast 2 Sequences results

PubMed

Entrez

BLAST

OMIM

Taxonomy

Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTN 2.1.2 [Nov-13-2000]

Match: ☐ Mismatch: ☐ gap open: ☐ gap extension: ☐x_dropoff: ☐ expect: ☐ wordsize: ☐ Filter ☐

Sequence 1 gi 4538106 Sequence 3 from Patent WO9740167. Length 3585 (2163 .. 5747)
 Sequence 2 gi 9630311 Gibbon leukemia virus, complete genome Length 8088 (1 .. 8088)

2 ☐ 1

NOTE: The statistics (bitscore and expect value) is calculated based on the size of nr database

NOTE: If protein translation is reversed, please repeat the search with reverse strand of the query sequence

Score = 798 bits (10671), Expect = 0.0
 Identities = 2598/3890 (66%), Gaps = 567/3890 (17%)
 Strand = Plus / Plus

Query: 2163 gggagacggggttcggacccccctccccgagccagggtaactttgaaggtgg
 REVERSE TRANSCRIPTASE 1 G R R G S D P L P E P R V T L K V
 194 gggagtcagggttcggacccccctccccgagccagggtaactttgaaggtgg
 Sbjet: 1 G S Q G S D P L P E P R V T L T V
 pol polyprotein

Query: 2222 accagttgagttcctggtgataccggagcggagcattcagtgctg-ctaca
 REVERSE TRANSCRIPTASE 21 P V E F L V D T G A E H S V L L Q
 254 ccca-ttgagttcctggtgacacccggagctgaacattcagttatgac-cca
 Sbjet: 21 P I E F L V D T G A E H S V L T Q
 pol polyprotein

Query: 2280 ggaaaacta----aaagaaaaaaatcctgg--gt-gatgggtgccacagg-
 REVERSE TRANSCRIPTASE 40 G K L K E K K S W V M G A T G
 2311 ggaaaagttagggtccagacgga---c--ggtcgtgga-aggagcgacaggc
 Sbjet: 40 G K V G S R R T V V E G A T G
 pol polyprotein

Query: 2331 -cagatccatggactaccggaagaacggtt---gacttggga-gtgggacg
 REVERSE TRANSCRIPTASE 57 Q Y P W T T R R T V D L G V G R
 2363 tc--taccctggaccacaaaaaga---cttttaaaatt-ggacat-aaaca
 Sbjet: 57 V Y P W T T K R L L K I G H K Q
 pol polyprotein

Query: 2386 actcgtttctggtcat-cctgagtgcccag-taccct-tctaggtagaga
 REVERSE TRANSCRIPTASE 75 H S F L V I P E C P V P L L G R D
 2417 actccttctggtcatccc-gagtgccctgct-cctctgt-tgggcagggg
 Sbjet: 75 H S F L V I P E C P A P L L G R D
 pol polyprotein

http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?0

22/02/01

Blast Result

```

Query:          2443 ccaaga-t---gggagctcaaaatttc---tttt--gaac-aaggaagaccag
REVERSE TRANSCRIPTASE 94  T K M G A Q I S F E Q G R P
                  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct:          2474 ccaa-actaaag---gcccgaga--tccagttttccg--ctgagg--gccca-
pol polyprotein 94  T K L K A Q I Q F S A E G P

```

```

Query:          2491 c-tgtgaataacaaaccatcactgtgttgacc--ctcc-aa--ttagatga
REVERSE TRANSCRIPTASE 110 S V N N K P I T V L T L Q L D D
                  ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
sbjct:          2522 catgggggagaac--gccct-actatg-tg-cctgggtcctaaacctggaaga
pol polyprotein 110 T W G E R P T M C L V L N L E E

```

```
Query:      2545 gactatattctccccaagtaaacctg-a---tcaagatat--a---cagtcc
REVERSE TRANSCRIPTASE 128 R L Y S P Q V K P D Q D I Q S
          |||||               || | ||||| | | ||   || | || |||
Sbjct:      2576 gacta-----catgaaaagccagtagcctc-ctctatcgacca-tcc
pol polyprotein 128 R L H E K P V P S S I D P S
```

```

Query:          2597 gcagtttcccca-agcctgggcagaaaccgcag--gg-atgggtttggcaaa
REVERSE TRANSCRIPTASE 146  Q F P Q A W A E T A G M G L A K
                  || ||||| | ||||| | || ||||| | |||||
                2625 gc-ttttccccactgtatgggcagaaa--g-agccggcatgggactagccaa
Sbjct:          145  L F P T V W A E R A G M G L A N
pol polyprotein

```

```

Query:          2653 cccacag-gttattca-a-ctgaaggcca-gtg-ctacaccagttacg-t
REVERSE TRANSCRIPTASE 164 P P Q V I Q L K A S A T P V S V
                  ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Sbjct:          2680 ccac-cagtggtagt-agagct-aagatcaggtgcct-caccagtggtgtt
pol polyprotein 163 P P V V V E L R S G A S P V A V

```

```

Query:          2707 acccccttgagtagagagggctcgagaaggaattggccgcatgttca-aagat
REVERSE TRANSCRIPTASE 182 Y P L S R E A R E G I W P H V Q R
                   || ||| | ||| |||| | |||| | | ||| | | ||| |
                2735 atccaatgagcaagaagctcggaaggtatcagacccacatccagaag-t
Sbjct:          181 Y P M S K E A R E G I R P H I Q K
pol polyprotein

```

```

Query:          2765 -ac--agggcatacctagt-tcctgtccastc-cccttggaatactccc-ctg
REVERSE TRANSCRIPTASE 202      Q G I L V P V Q S P W N T P L
                || |||| || || ||||| | || ||||| ||| |||
                2791 gacctagggt-gtcttggtgcctgt-cggtcgccc-tggaatac-ccctctg
sbjct:          200      D L G V L V P C R S P W N T P L
pol polyprotein

```

```

Query:          2819 caggaagcctgggaccaatgattatcgaccagtacaggacttgagagaggtc
REVERSE TRANSCRIPTASE 220  R K P G T N D Y R P V Q D L R E V
                   | | | | | | | | | | | | | | | | | | | | | | | |
                   2647 aaaaaagccagggaacaaatgactatcggaaggttcaagacctgagagaaatt
Shjct:          219  K K P G T N D Y R P V Q D L R E I
pol polyprotein

```

```

Query:          2879 ggtgcaggacataca-cccaacggtcccgaaaccttataa-cctcttgagcg
REVERSE TRANSCRIPTASE 240  V Q D I H P T V P N P Y N L L S
                  ||| |||| || || || || || || || || || || || || || || ||
Sbjct:          2907 ggtacaggatattatcc-accagtcctcaaaccttacaattcttc-gagtt
pol polyprotein 239  V Q D I H P T V P N P Y N L L S

```

```

Query:      2937 cctgaacgga-ac-tggtacacagtagtattgga-cttaaaagatgaccttttct
REVERSE     259  P E R N W Y T V L D L K D A F F
              ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
              2065 cct-aggc-tatacttggtactcagtccttagatc-tcaaggatgaccttttct
Subject:
http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?0
22/02/01

```

Blast Result

Page 1 of 5

Attachment



Blast 2 Sequences results

PubMed

Entrez

BLAST

OMIM

Taxonomy

Structure

BLAST 2 SEQUENCES RESULTS VERSION BLASTN 2.1.2 [Nov-13-2000]

Match: 5 Mismatch: 2 gap open: 1 gap extension: 1

x dropoff: 100 expect: 10.00 wordsize: 7 Filter ☐

Sequence 1 gi_4538106 Sequence 3 from Patent WO9740167. Length 1971 (5620 .. 7590)

Sequence 2 gi_963031.1 Gibbon leukemia virus, complete genome Length 2449 (5552.. 8000)

NOTE:The statistics (bitscore and expect value) is calculated based on the size of nr database

NOTE: If protein translation is reversed, please repeat the search with reverse strand of the query sequence

Score = 410 bits (5448), Expect = e-112
Identities = 1396/2290 (60%), Gaps = 560/2290 (24%)
Strand = Plus / Plus

```

Query:                5620 atgcatccacgttaagcoggogccacctccog-actc-g--ggg---tggg
VIRAL ENVELOPE PROTEIN 1   M H P T L S R R H L P T R G G
REVERSE TRANSCRIPTASE 1153 H A S H V K P A P P P D S G W
                        ||||| || ||| || || || || || || || || || || || || || ||
Sbjct:                5573 atgcttctcacctcaaacctgcaccaccttcggcac-cagatgagtcctgg-
env protein           8   M L L T S N L K H H L R H Q M S P G
pol polyprotein      1127 H A S H L K P A P P S A P D E S W

```

```

Query:                5672 aaagactgagaatccccttaagcttgcctcca-tc-gc--ctggttcccta
VIRAL ENVELOPE PROTEIN 18  K R L R I P L S F A S I A W F L
REVERSE TRANSCRIPTASE 1171  K T E N P L K L R L H R L V P Y
                        ||||| ||| ||||| || | | | | | | | | |
Sbjct:                5631 aaagactgacatcctctttaagct--gcgt--attcgcgggc-gg--cgggg
env protein           27  K R L I I L L S C V F G G G G
pol polyprotein       1147 K T D H P L K L R I R R R R D

```

```

Query:          5724 -aacaat---aactcc--ccag-----g-c--ca-----gta--gt--
VIRAL ENVELOPE PROTEIN 36      T I      T P      Q      A      S      S
REVERSE TRANSCRIPTASE 1188      N N      N S      P      G      Q      ^  ^
                               || |||  ||| ||  ||||  ||  ||  ||| ||
Sbjct:          5684 caa-aataagaaacccccaccagcccatgacccctcacttgcaggtactotcc
env protein     45      Q  N K N P H Q P M T L T W Q V L S
pol polypeptin 1164 A      K ^^^

```

```

Query:          5754  ---cett-----at---agacagctcgaaaccccatagacctt--tato
VIRAL ENVELOPE PROTEIN 46      L      I      D S S N P H R P      L S
                        |  ||      ||  || || | || | || | || |
Sbjct:          5743  agacgtgtctgtgggacacaaaggcag-tccagccccccttgga-cttggtggc
env protein      65      D V V W D T K A V Q P P W T W W

```

<http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi?0>

22/02/01

Page 2 of 5

```

Query:                6230 atgagctggggaat-agttttttattatatg-g-cggggga-gcagg-gtcc
VIRAL ENVELOPE PROTEIN 204 Y E L G N S F L L Y G G G A G S
                        | | | | | | | | | | | | | | | | | |
Sbjct:                6254 aaaacctggggattaag-----attctatgtgtctggacatccaggcgtaac
env protein           235  K T W G L R               F Y V S G H P G V

```

22/02/01